

CLAIMS

What is claimed is:

- 1 1. A machine-implemented method for estimating how a particular user would rate a
2 particular item from a plurality of items, the method comprising the machine-
3 implemented steps of:
4 identifying, from the plurality of items, one or more other items that have ratings
5 similar to the particular item;
6 selecting a set of other users from a plurality of other users based upon user
7 selection criteria;
8 identifying one or more other users from the set of other users that have given
9 ratings to a subset of the one or more other items, wherein the ratings are
10 statistically similar to ratings given by the particular user to the subset of
11 the one or more other items; and
12 generating an estimation of how the particular user would rate the particular item
13 based upon ratings for the particular item given by the one or more other
14 users.
- 1 2. The machine implemented method as recited in Claim 1, wherein the step of
2 identifying, from the plurality of items, one or more other items that have ratings
3 similar to the particular item is performed off-line in advance of the other steps.
- 1 3. A machine-readable medium for estimating how a particular user would rate a
2 particular item from a plurality of items, the machine-readable medium carrying

3 instructions which, when processed by one or more processors, cause
4 performance of the steps of:
5 identifying, from the plurality of items, one or more other items that have ratings
6 similar to the particular item;
7 selecting a set of other users from a plurality of other users based upon user
8 selection criteria;
9 identifying one or more other users from the set of other users that have given
10 ratings to a subset of the one or more other items, wherein the ratings are
11 statistically similar to ratings given by the particular user to the subset of
12 the one or more other items; and
13 generating an estimation of how the particular user would rate the particular item
14 based upon ratings for the particular item given by the one or more other
15 users.

1 4. The machine-readable medium as recited in Claim 3, wherein the step of
2 identifying, from the plurality of items, one or more other items that have ratings
3 similar to the particular item is performed off-line in advance of the other steps.

1 5. An apparatus for estimating how a particular user would rate a particular item
2 from a plurality of items, the apparatus comprising a memory storing instructions
3 which, when processed by one or more processors, cause performance of the steps
4 of:
5 identifying, from the plurality of items, one or more other items that have ratings
6 similar to the particular item;

7 selecting a set of other users from a plurality of other users based upon user
8 selection criteria;
9 identifying one or more other users from the set of other users that have given
10 ratings to a subset of the one or more other items, wherein the ratings are
11 statistically similar to ratings given by the particular user to the subset of
12 the one or more other items; and
13 generating an estimation of how the particular user would rate the particular item
14 based upon ratings for the particular item given by the one or more other
15 users.

1 6. The apparatus as recited in Claim 5, wherein the step of identifying, from the
2 plurality of items, one or more other items that have ratings similar to the particular
3 item is performed off-line in advance of the other steps.

1 7. A machine-implemented method for estimating how a particular user would rate a
2 particular item from a plurality of items, the method comprising the machine-
3 implemented steps of:
4 identifying one or more items from the plurality of items that have ratings similar
5 to the particular item;
6 identifying one or more other users that have given ratings to the one or more
7 items that are substantially similar to ratings given by the particular user to
8 the one or more items; and

9 generating an estimation of how the particular user would rate the particular item
10 based upon ratings for the particular item given by the one or more other
11 users.

1 8. The machine implemented method as recited in Claim 7, wherein the step of
2 identifying, from the plurality of items, one or more other items that have ratings
3 similar to the particular item is performed off-line in advance of the other steps.
4

1 9. The machine-implemented method as recited in Claim 7, wherein identifying one
2 or more items from a plurality of items that have ratings similar to the particular
3 item includes calculating a statistical correlation between the particular item and
4 the other items from the plurality of items.

1 10. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented step of determining whether the particular item and each of
3 the one or more items have at least a specified number of co-ratings.

1 11. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented step of determining whether the particular item and one of
3 the one or more items have at least a specified fraction of co-ratings between the
4 particular item and the one of the one or more items divided by the number of
5 ratings of either the particular item or the one of the one or more items.

1 12. The machine-implemented method as recited in Claim 7, wherein the step of
2 identifying one or more items from the plurality of items that have ratings similar to
3 the particular item includes identifying whether the particular item and the one or
4 more items have a number of co-raters greater than a specified fraction of a
5 statistically expected number, given the number of ratings of each item counted
6 separately.

1 13. The machine-implemented method as recited in Claim 7, wherein the step of
2 identifying one or more items from the plurality of items is performed by one or
3 more external agents.

1 14. The machine-implemented method as recited in Claim 7, wherein the step of
2 identifying one or more items from the plurality of items that have ratings similar to
3 the particular item includes consideration of one or more of a statistical correlation
4 between the particular item and the one or more items, whether the particular item
5 and the one or more items have at least a specified number of co-ratings, whether the
6 particular item and the one or more items have at least a specified fraction of co-
7 ratings, and whether the particular item and the one or more items have at least a
8 specified fraction of the statistically expected number of co-ratings.

1 15. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented step of selecting the one or more other users from a plurality
3 of other users based upon each of the one or more other users having at least a
4 specified minimum number of ratings.

1 16. The machine-implemented method as recited in Claim 7, further comprising the

2 machine-implemented step of selecting the one or more other users from a plurality
3 of other users based upon each of the one or more other users having less than a
4 specified maximum number of ratings.

1 17. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented step of selecting the one or more other users from a plurality
3 of other users based upon each of the one or more other users having an average
4 rating within a specified range.

1 18. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented step of selecting the one or more other users from a plurality
3 of other users based upon each of the one or more other users having a rating
4 distribution within a specified tolerance of a Normal distribution.

1 19. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented step of selecting the one or more other users from a plurality
3 of other users based upon each of the one or more other users having a rating
4 standard deviation within a specified range.

1 20. The machine-implemented method as recited in Claim 7, wherein identifying one or
2 more other users that have given ratings to the one or more items that are
3 substantially similar to ratings given by the particular user to the one or more items
4 includes consideration of one or more of whether the one or more other users have at
5 least a specified minimum number of co-ratings, whether the one or more other
6 users have less than a specified maximum number of ratings, whether the one or
7 more other users have an average rating within a specified range, whether the one or

8 more other users have a rating distribution within a specified tolerance of a Normal
9 distribution, and whether the one or more other users have a rating standard
10 deviation within a specified range.

1 21. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing the one or more other users for whom a
3 statistical correlation with the particular user of co-ratings of the one or more items
4 is greater than a specified threshold.

1 22. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing the one or more other users for whom a
3 vector cosine distance of co-ratings of the one or more items is smaller than a
4 specified threshold.

1 23. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing the one or more other users who have co-
3 rated a subset of the one or more items exactly the same as the particular user.

1 24. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing the one or more other users who have
3 rated co-rated a subset of the one or more items within a specified threshold of the
4 ratings by the particular user.

1 25. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes consideration of one or more of whether the one or
3 more users have a correlation of co-ratings of the one or more items greater than a

4 specified threshold, whether the one or more users have a vector cosine distance of
5 co-ratings of the one or more items smaller than a specified threshold, whether the
6 one or more users have co-rated a subset of the one or more items exactly the same,
7 and whether the one or more users having co-rated a subset of the one or more items
8 within a specified threshold of the ratings by the particular user.

1 26. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing a specified number of other users most
3 similar to the particular user.

1 27. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing other users whose similarity to the
3 particular user is better than a specified threshold.

1 28. The machine-implemented method as recited in Claim 7, wherein identifying the
2 one or more other users includes choosing up to a specified number of other users
3 most similar to the particular user, where all the chosen other users have similarity
4 better than a specified threshold.

1 29. The machine-implemented method as recited in Claim 7, wherein generating an
2 estimation of how the user would rate the particular item based upon ratings for
3 the particular item given by the one or more other users includes determining an
4 average rating of the particular item by the one or more other users.

1 30. The machine-implemented method as recited in Claim 7, wherein generating an
2 estimation of how the particular user would rate the particular item based upon
3 ratings for the particular item given by the one or more other users includes
4 determining an average rating of the particular item by the one or more other users.

1 31. The machine-implemented method as recited in Claim 7, further comprising the
2 machine-implemented steps of:
3 generating, for each of the one or more users, a separate estimation of how the
4 particular user would rate the particular item based upon the rating for the
5 particular item given by each of the one or more other users, and
6 calculating a confidence metric based upon disparities between the separate
7 estimations.

1 32. The machine-implemented method as recited in Claim 7, wherein the plurality of
2 items is a plurality of movies, the one or more items is one or more movies and the
3 particular item is a particular movie.

1 33. The machine-implemented method as recited in Claim 7, wherein the plurality of
2 items is a plurality of music items, the one or more items is one or more music items
3 and the particular item is a particular music item.

1 34. A machine-readable medium for estimating how a particular user would rate a
2 particular item from a plurality of items, the machine-readable medium carrying
3 instructions which, when processed by one or more processors, cause
4 performance of the steps of:

5 identifying, from the plurality of items, one or more items that have ratings
6 similar to the particular item;
7 identifying one or more other users that have given ratings to the one or more
8 items that are substantially similar to ratings given by the particular user to
9 the one or more items; and
10 generating an estimation of how the particular user would rate the particular item
11 based upon ratings for the particular item given by the one or more other
12 users.

1 35. The machine-readable medium as recited in Claim 34, wherein the step of
2 identifying, from the plurality of items, one or more other items that have ratings
3 similar to the particular item is performed off-line in advance of the other steps.

1 36. The machine-readable medium as recited in Claim 34, wherein identifying one or
2 more items from a plurality of items that have ratings similar to the particular item
3 includes calculating a statistical correlation between the particular item and the
4 other items from the plurality of items.

1 37. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more
3 processors, cause the one or more processors to perform the additional step of
4 determining whether the particular item and each of the one or more items have at
5 least a specified number of co-ratings.

1 38. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more

3 processors, cause the one or more processors to perform the additional step of
4 determining whether the particular item and one of the one or more items have at
5 least a specified fraction of co-ratings between the particular item and the one of
6 the one or more items divided by the number of ratings of either the particular
7 item or the one of the one or more items.

1 39. The machine-readable medium as recited in Claim 34, wherein the step of
2 identifying one or more items from the plurality of items that have ratings similar to
3 the particular item includes identifying whether the particular item and the one or
4 more items have a number of co-raters greater than a specified fraction of a
5 statistically expected number, given the number of ratings of each item counted
6 separately.

1 40. The machine-readable medium as recited in Claim 34, wherein the step of
2 identifying one or more items from the plurality of items is performed by one or
3 more external agents.

1 41. The machine-readable medium as recited in Claim 34, wherein the step of
2 identifying one or more items from the plurality of items that have ratings similar to
3 the particular item includes consideration of one or more of a statistical correlation
4 between the particular item and the one or more items, whether the particular item
5 and the one or more items have at least a specified number of co-ratings, whether the
6 particular item and the one or more items have at least a specified fraction of co-
7 ratings, and whether the particular item and the one or more items have at least a
8 specified fraction of the statistically expected number of co-ratings.

1 42. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having at least a minimum number of ratings.

1 43. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having less than a maximum number of ratings.

1 44. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having an average rating within a specified range.

1 45. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having a rating distribution within a specified tolerance of a
6 Normal distribution.

1 46. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having a rating standard deviation within a specified range.

1 47. The machine-implemented method as recited in Claim 34, wherein identifying one
2 or more other users that have given ratings to the one or more items that are
3 substantially similar to ratings given by the particular user to the one or more items
4 includes consideration of one or more of whether the one or more other users have at
5 least a specified minimum number of co-ratings, whether the one or more other
6 users have less than a specified maximum number of ratings, whether the one or
7 more other users have an average rating within a specified range, whether the one or
8 more other users have a rating distribution within a specified tolerance of a Normal
9 distribution, and whether the one or more other users have a rating standard
10 deviation within a specified range.

1 48. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing the one or more other users for whom a
3 statistical correlation with the particular user of co-ratings of the one or more items
4 is greater than a specified threshold.

1 49. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing the one or more other users for whom a
3 vector cosine distance of co-ratings of the one or more items is smaller than a

4 specified threshold.

1 50. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing the one or more other users who have co-
3 rated a subset of the one or more items exactly the same as the particular user.

1 51. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing the one or more other users who have
3 rated co-rated a subset of the one or more items within a specified threshold of the
4 ratings by the particular user.

1 52. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes consideration of one or more of whether the one or
3 more users have a correlation of co-ratings of the one or more items greater than a
4 specified threshold, whether the one or more users have a vector cosine distance of
5 co-ratings of the one or more items smaller than a specified threshold, whether the
6 one or more users have co-rated a subset of the one or more items exactly the same,
7 and whether the one or more users having co-rated a subset of the one or more items
8 within a specified threshold of the ratings by the particular user.

1 53. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing a specified number of other users most
3 similar to the particular user.

1 54. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing other users whose similarity to the

3 particular user is better than a specified threshold.

1 55. The machine-implemented method as recited in Claim 34, wherein identifying the
2 one or more other users includes choosing up to a specified number of other users
3 most similar to the particular user, where all the chosen other users have similarity
4 better than a specified threshold.

1 56. The machine-readable medium as recited in Claim 34, wherein generating an
2 estimation of how the particular user would rate the particular item based upon
3 ratings for the particular item given by the one or more other users includes
4 determining an average rating of the particular item by the one or more other
5 users.

1 57. The machine-readable medium as recited in Claim 34, wherein generating an
2 estimation of how the particular user would rate the particular item based upon
3 ratings for the particular item given by the one or more other users includes
4 determining an average rating of the particular item by the one or more other users.

1 58. The machine-readable medium as recited in Claim 34, further comprising one or
2 more additional instructions which, when processed by the one or more
3 processors, cause the one or more processors to perform the additional steps of:
4 generating, for each of the one or more users, a separate estimation of how the
5 particular user would rate the particular item based upon the rating for the
6 particular item given by each of the one or more other users, and
7 calculating a confidence metric based upon disparities between the separate
8 estimations.

1 59. The machine-readable medium as recited in Claim 34, wherein the plurality of items
2 is a plurality of movies, the one or more items is one or more movies and the
3 particular item is a particular movie.

1 60. The machine-readable medium as recited in Claim 34, wherein the plurality of items
2 is a plurality of music items, the one or more items is one or more music items and
3 the particular item is a particular music item.

1 61. An apparatus for estimating how a particular user would rate a particular item
2 from a plurality of items, the apparatus comprising a memory storing instructions
3 which, when processed by one or more processors, cause the one or more
4 processors to perform the steps of:
5 identifying, from the plurality of items, one or more items that have ratings
6 similar to the particular item;
7 identifying one or more other users that have given ratings to the one or more
8 items that are substantially similar to ratings given by the particular user to
9 the one or more items; and
10 generating an estimation of how the particular user would rate the particular item
11 based upon ratings for the particular item given by the one or more other
12 users.

1 62. The apparatus as recited in Claim 61, wherein the step of identifying, from the
2 plurality of items, one or more other items that have ratings similar to the particular
3 item is performed off-line in advance of the other steps.

1 63. The apparatus as recited in Claim 61, wherein identifying one or more items from
2 a plurality of items that have ratings similar to the particular item includes
3 calculating a statistical correlation between the particular item and the other items
4 from the plurality of items.

1 64. The apparatus as recited in Claim 61, wherein the memory further comprises one
2 or more additional instructions which, when processed by the one or more
3 processors, cause the one or more processors to perform the additional step of
4 determining whether the particular item and each of the one or more items have at
5 least a specified number of co-ratings.

1 65. The apparatus as recited in Claim 61, wherein the memory further comprises one
2 or more additional instructions which, when processed by the one or more
3 processors, cause the one or more processors to perform the additional step of
4 determining whether the particular item and one of the one or more items have at
5 least a specified fraction of co-ratings between the particular item and the one of
6 the one or more items divided by the number of ratings of either the particular
7 item or the one of the one or more items.

1 66. The apparatus as recited in Claim 61, wherein the step of identifying one or more
2 items from the plurality of items that have ratings similar to the particular item
3 includes identifying whether the particular item and the one or more items have a
4 number of co-raters greater than a specified fraction of a statistically expected
5 number, given the number of ratings of each item counted separately.

1 67. The apparatus as recited in Claim 61, wherein the step of identifying one or more
2 items from the plurality of items is performed by one or more external agents.

1 68. The apparatus as recited in Claim 61, wherein the step of identifying one or more
2 items from the plurality of items that have ratings similar to the particular item
3 includes consideration of one or more of a statistical correlation between the
4 particular item and the one or more items, whether the particular item and the one or
5 more items have at least a specified number of co-ratings, whether the particular
6 item and the one or more items have at least a specified fraction of co-ratings, and
7 whether the particular item and the one or more items have at least a specified
8 fraction of the statistically expected number of co-ratings.

1 69. The apparatus as recited in Claim 61, wherein the memory further comprises one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having at least a minimum number of ratings.

1 70. The apparatus as recited in Claim 61, wherein the memory further comprises one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one
4 or more other users from a plurality of other users based upon each of the one or
5 more other users having less than a maximum number of ratings.

1 71. The apparatus as recited in Claim 61, wherein the memory further comprises one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one or
4 more other users from a plurality of other users based upon each of the one or more
5 other users having an average rating within a specified range.

1 72. The apparatus as recited in Claim 61, wherein the memory further comprises one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one or
4 more other users from a plurality of other users based upon each of the one or more
5 other users having a rating distribution within a specified tolerance of a Normal
6 distribution.

1 73. The apparatus as recited in Claim 61, wherein the memory further comprises one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional step of selecting the one or
4 more other users from a plurality of other users based upon each of the one or more
5 other users having a rating standard deviation within a specified range.

1 74. The apparatus as recited in Claim 61, wherein identifying one or more other users that
2 have given ratings to the one or more items that are substantially similar to ratings
3 given by the particular user to the one or more items includes consideration of one or
4 more of whether the one or more other users have at least a specified minimum
5 number of co-ratings, whether the one or more other users have less than a specified
6 maximum number of ratings, whether the one or more other users have an average

7 rating within a specified range, whether the one or more other users have a rating
8 distribution within a specified tolerance of a Normal distribution, and whether the one
9 or more other users have a rating standard deviation within a specified range.

1 75. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing the one or more other users for whom a statistical correlation with
3 the particular user of co-ratings of the one or more items is greater than a specified
4 threshold.

1 76. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing the one or more other users for whom a vector cosine distance of
3 co-ratings of the one or more items is smaller than a specified threshold.

1 77. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing the one or more other users who have co-rated a subset of the one
3 or more items exactly the same as the particular user.

1 78. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing the one or more other users who have rated co-rated a subset of the
3 one or more items within a specified threshold of the ratings by the particular user.

1 79. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes consideration of one or more of whether the one or more users have a
3 correlation of co-ratings of the one or more items greater than a specified threshold,
4 whether the one or more users have a vector cosine distance of co-ratings of the one or
5 more items smaller than a specified threshold, whether the one or more users have co-

6 rated a subset of the one or more items exactly the same, and whether the one or more
7 users having co-rated a subset of the one or more items within a specified threshold of
8 the ratings by the particular user.

1 80. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing a specified number of other users most similar to the particular user.

1 81. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing other users whose similarity to the particular user is better than a
3 specified threshold.

1 82. The apparatus as recited in Claim 61, wherein identifying the one or more other users
2 includes choosing up to a specified number of other users most similar to the
3 particular user, where all the chosen other users have similarity better than a specified
4 threshold.

1 83. The apparatus as recited in Claim 61, wherein generating an estimation of how the
2 particular user would rate the particular item based upon ratings for the particular
3 item given by the one or more other users includes determining an average rating of
4 the particular item by the one or more other users.

1 84. The apparatus as recited in Claim 61, wherein generating an estimation of how the
2 particular user would rate the particular item based upon ratings for the particular item
3 given by the one or more other users includes determining an average rating of the
4 particular item by the one or more other users.

1 85. The apparatus as recited in Claim 61, wherein the memory further comprises one or
2 more additional instructions which, when processed by the one or more processors,
3 cause the one or more processors to perform the additional steps of:
4 generating, for each of the one or more users, a separate estimation of how the
5 particular user would rate the particular item based upon the rating for the
6 particular item given by each of the one or more other users, and
7 calculating a confidence metric based upon disparities between the separate
8 estimations.

1 86. The apparatus as recited in Claim 61, wherein the plurality of items is a plurality of
2 movies, the one or more items is one or more movies and the particular item is a
3 particular movie.

1 87. The apparatus as recited in Claim 61, wherein the plurality of items is a plurality of
2 music items, the one or more items is one or more music items and the particular item
3 is a particular music item.

1 88. An apparatus configured to:
2 generate an estimation of how a particular user would rate a particular item from a
3 plurality of items by:
4 identifying, from the plurality of items, one or more items that have ratings
5 similar to the particular item;
6 identifying one or more other users that have given ratings to the one or
7 more items that are substantially similar to ratings given by the
8 particular user to the one or more items; and

9 generating the estimation of how the particular user would rate the
10 particular item based upon ratings for the particular item given by
11 the one or more other users.

1 89. The apparatus as recited in Claim 88, wherein the step of generating the estimation
2 is performed in advance of the need for the estimation and data that indicates the
3 estimation is stored in a memory.

1 90. The apparatus as recited in Claim 89, wherein the apparatus is further configured to
2 write data that indicates the estimation to a non-volatile storage.

1 91. The apparatus as recited in Claim 90, wherein the non-volatile storage is one or
2 more magnetic or solid-state disc systems, all or portions of which are memory-
3 mapped into a random-access memory as needed.

1 92. The apparatus as recited in Claim 90, wherein the apparatus is further configured to
2 retrieve the data that indicates the estimation from the non-volatile storage and
3 write the data into a volatile memory.

1 93. The apparatus as recited in Claim 88, wherein the apparatus is further configured
2 to in response to one or more ratings being updated or added for any of the one or
3 more other items or the particular item:
4 identifying one or more other items from the plurality of items that have ratings
5 similar to the particular item, wherein the one or more other items may
6 include any of the one or more items;

7 identifying one or more users that have given ratings to the one or more other
8 items that are substantially similar to ratings given by the particular user to
9 the one or more other items;
10 generating a revised estimation of how the particular user would rate the
11 particular item based upon ratings for the particular item given by the one
12 or more users; and
13 storing data that indicates the revised estimation of how the particular user would
14 rate the particular item.

1 94. The apparatus as recited in Claim 88, wherein the apparatus is further configured
2 to in response to one or more ratings being updated or added for any of the one or
3 more other items or the particular item:
4 identifying one or more estimated ratings that are affected by the one or more
5 ratings being updated or added;
6 generating one or more updated estimated ratings for the one or more estimated
7 ratings that are affected by the one or more ratings being updated or
8 added; and
9 storing data that indicates the one or more updated estimated ratings.

1
1 95. The apparatus as recited in Claim 88, wherein the apparatus is further configured
2 to periodically generate and update similarity data that identifies similarities
3 between items from the plurality of items based upon ratings associated with the
4 plurality of items.

1

- 1 96. The apparatus as recited in Claim 88, wherein the apparatus is further configured
2 to select the one or more other users from a plurality of users based upon selection
3 criteria that includes whether a particular user has a specified number of ratings
4 and to generate and store data that indicates the selected one or more other users.
- 1 97. The apparatus as recited in Claim 88, further comprising one or more prioritized
2 task queues from which tasks are assigned to be processed based upon a priority
3 associated with each task and wherein tasks associated with updating estimated
4 ratings for active users are assigned the highest priority.
- 1 98. The apparatus as recited in Claim 88, wherein the plurality of items is a plurality
2 of movies, the one or more items is one or more movies and the particular item is
3 a particular movie.
- 1 99. The apparatus as recited in Claim 88, wherein the plurality of items is a plurality of
2 music items, the one or more items is one or more music items and the particular
3 item is a particular music item.
- 1 100. An apparatus comprising:
2 a plurality of processes assigned to generate a plurality of estimated ratings of
3 how users would rate items from a plurality of items, wherein generating
4 an estimate of how a particular user would rate a particular item from the
5 plurality of items is performed by:
6 identifying one or more items from the plurality of items that have ratings
7 similar to the particular item;

8 identifying one or more other users that have given ratings to the one or
9 more items that are substantially similar to ratings given by the
10 particular user to the one or more items; and
11 generating the estimation of how the particular user would rate the
12 particular item based upon ratings for the particular item given by
13 the one or more other users.

1 101. The apparatus as recited in Claim 100, further comprising one or more processes
2 that are configured to generate and store ratings data that indicates estimated
3 ratings, wherein each such process is not given access to ratings data generated
4 and stored by other such processes.

5 102. The apparatus as recited in Claim 101, further comprising one or more processes
6 assigned to select reference users and identify similarities between items from the
7 plurality of items based upon ratings associated with the plurality of items and to
8 generate and store similarity data that indicates the identified similarities between
9 items from the plurality of items, wherein the one or more processes have
10 read/write access to the similarity and reference user data and the one or more
11 other processes have read-only access to the similarity data.

1 103. The apparatus as recited in Claim 102, wherein the similarity and reference user
2 data is copied periodically and the one or more recommendation processes have
3 access to the copied similarity and reference user data.

1 104. The apparatus as recited in Claim 100, further comprising an additional process
2 assigned to select the one or more other users from a plurality of users based upon

3 selection criteria that includes whether a user has a specified number of ratings
4 and to generate data that indicates the selected one or more other users, wherein
5 the additional process has read/write access to the data and the plurality of
6 processes have read-only access to the data.

1 105. The apparatus as recited in Claim 100, wherein the plurality of items is a plurality of
2 movies, the one or more items is one or more movies and the particular item is a
3 particular movie.

1 106. The apparatus as recited in Claim 100, wherein the plurality of items is a plurality of
2 music items, the one or more items is one or more music items and the particular
3 item is a particular music item.

1 107. A machine-implemented method for estimating how a particular user would rate a
2 particular item from a plurality of items, the method comprising the machine-
3 implemented steps of:
4 identifying one or more items from the plurality of items that are similar to the
5 particular item by associating semantic feature values with the plurality of
6 items and the particular item, where each feature type has a match-score
7 and a mismatch-score, either of which may be positive, zero, or negative,
8 and scoring the similarity of items is performed by
9 finding the features that have matching values and summing the
10 associated match-score, and
11 finding the features that have mismatched values and summing the
12 associated mismatch-score,

13 wherein items with a summed similarity score above a specified
14 threshold are considered similar or the specified number of
15 items with the highest similarity score are considered
16 similar;
17 identifying one or more other users that have given ratings to the one or more
18 items that are substantially similar to ratings given by the particular user to
19 the one or more items; and
20 generating an estimation of how the particular user would rate the particular item
21 based upon ratings for the particular item given by the one or more other
22 users.

1 108. A machine-readable medium for estimating how a particular user would rate a
2 particular item from a plurality of items, the machine-readable medium carrying
3 instructions which, when processed by one or more processors, cause
4 performance of the steps of:
5 identifying one or more items from the plurality of items that are similar to the
6 particular item by associating semantic feature values with the plurality of
7 items and the particular item, where each feature type has a match-score
8 and a mismatch-score, either of which may be positive, zero, or negative,
9 and scoring the similarity of items is performed by
10 finding the features that have matching values and summing the
11 associated match-score, and
12 finding the features that have mismatched values and summing the
13 associated mismatch-score,

14 wherein items with a summed similarity score above a specified
15 threshold are considered similar or the specified number of
16 items with the highest similarity score are considered
17 similar;
18 identifying one or more other users that have given ratings to the one or more
19 items that are substantially similar to ratings given by the particular user to
20 the one or more items; and
21 generating an estimation of how the particular user would rate the particular item
22 based upon ratings for the particular item given by the one or more other
23 users.

1 109. An apparatus for estimating how a particular user would rate a particular item
2 from a plurality of items, the apparatus comprising a memory storing instructions
3 which, when processed by one or more processors, cause performance of the
4 machine-implemented steps of:
5 identifying one or more items from the plurality of items that are similar to the
6 particular item by associating semantic feature values with the plurality of
7 items and the particular item, where each feature type has a match-score
8 and a mismatch-score, either of which may be positive, zero, or negative,
9 and scoring the similarity of items is performed by
10 finding the features that have matching values and summing the
11 associated match-score, and
12 finding the features that have mismatched values and summing the
13 associated mismatch-score,

14 wherein items with a summed similarity score above a specified
15 threshold are considered similar or the specified number of
16 items with the highest similarity score are considered
17 similar;
18 identifying one or more other users that have given ratings to the one or more
19 items that are substantially similar to ratings given by the particular user to
20 the one or more items; and
21 generating an estimation of how the particular user would rate the particular item
22 based upon ratings for the particular item given by the one or more other
23 users.